Mother's own milk compared with formula milk for feeding preterm or low birth weight infants: Systematic review and meta-analysis

Natalie A. Strobel
Edith Cowan University, n.strobel@ecu.edu.au

Claire Adams

Daniel R. McAullay
Edith Cowan University, d.mcaullay@ecu.edu.au

Karen M. Edmond

Follow this and additional works at: https://ro.ecu.edu.au/ecuworks2022-2026

Part of the Maternal, Child Health and Neonatal Nursing Commons

10.1542/peds.2022-057092D

This Journal Article is posted at Research Online.
https://ro.ecu.edu.au/ecuworks2022-2026/1051
Mother’s Own Milk Compared With Formula Milk for Feeding Preterm or Low Birth Weight Infants: Systematic Review and Meta-analysis

Natalie A. Strobel, PhD,a Claire Adams, PhD,a Daniel R. McAullay, PhD,a Karen M. Edmond, MBBS, PhDb

OBJECTIVES: We assessed the effect of feeding preterm or low birth weight infants with infant formula compared with mother’s own milk on mortality, morbidity, growth, neurodevelopment, and disability.

METHODS: We searched Medline (Ovid), Embase (Ovid), and Cochrane Central Register of Controlled Studies to October 1, 2021.

RESULTS: Forty-two studies enrolling 89,638 infants fulfilled the inclusion criteria. We did not find evidence of an effect on mortality (odds ratio [OR] 1.26, 95% confidence interval [CI] 0.91–1.76), infection (OR 1.52, 95% CI 0.98–2.37), cognitive neurodevelopment (standardized mean difference —1.30, 95% CI —3.53 to 0.93), or on growth parameters. Formula milk feeding increased the risk of necrotizing enterocolitis (OR 2.99, 95% CI 1.75–5.11). The Grading of Recommendations Assessment, Development, and Evaluation certainty of evidence was low for mortality and necrotizing enterocolitis, and very low for neurodevelopment and growth outcomes.

CONCLUSIONS: In preterm and low birth weight infants, low to very low-certainty evidence indicates that feeding with infant formula compared with mother’s own milk has little effect on all-cause mortality, infection, growth, or neurodevelopment, and a higher risk of developing necrotizing enterocolitis.
Preterm (<37 weeks’ gestation) and low birth weight (LBW) (<2.5 kg) infants have limited nutrient reserves at birth and are subject to many physiologic and metabolic stresses that increase their nutrient needs. Formula milks (eg, artificial infant formulas) can be manipulated to contain higher amounts of nutrients (such as protein) than mother’s own milk. However, formula milks do not contain the immunomodulators and nutrients present in human milk that stimulate the immune system, protect the immature gut, and promote neurodevelopment.

There are many new infant formulas; however, the last systematic reviews were conducted in 2011 and 2019 and there have been no recent reviews of the effectiveness of infant formula and other formula milks compared with mother’s own milk on outcomes in preterm and LBW infants.

Our primary objective was to evaluate the effect of formula milks compared with mother’s own milk on primary outcomes (mortality, morbidity, growth, neurodevelopment, and disability) in preterm and LBW infants. Our secondary objectives were to determine the effect of gestational age (<32 weeks), birth weight (<1.5 kg), and exclusivity of mother’s own milk (ie, if the mother’s own milk was the sole diet [ie, 100% mother’s own milk]), on health outcomes of preterm or LBW infants.

METHODS

This review was registered in PROSPERO (#CRD42021283008). Preferred Reporting Items for Systematic Reviews and Meta-Analyses-Protocol guidance was followed.

Type and Setting of Included Studies

We included randomized controlled studies (RCTs), cohort, cross-sectional, and case-control studies. Case reports and studies published in abstract form only were excluded. All settings were included, such as home and health facility, within any country.

Participants

Only preterm or LBW infants were included. Normal weight or term infants were excluded.

Intervention and Control Groups

The intervention was any formula milk which included artificial infant formula (including cow’s milk protein, soy protein, other protein, or hydrolyzed formula) or other animal milk. Infants could receive mother’s own milk or donor human milk, as long as most (>50%) of the milk was formula. Infants had to receive the intervention in the neonatal period (0–27 days).

The control was mother’s own milk. Infants could also receive other milks (ie, donor human milk, formula milk, formula) as long as most (>50%) of the milk was mother’s own.

In both the intervention and control groups, the infants could receive any water-based fluids.

Outcomes

The primary outcome was infant all-cause mortality.

The secondary outcomes were:

- necrotizing enterocolitis as defined by the study authors;
- severe infections (eg, sepsis, pneumonia, meningitis) as defined by the study authors;
- neurodevelopment defined as: neurodevelopmental scores measured using validated assessment tools in the main domains (cognitive, motor, language) of standardized assessment tools such as the Bayley Scales of Infant and Toddler Development, Third Edition, or the Weschler Wechsler Intelligence Scale for Children;
- disability defined as: nonambulant cerebral palsy; developmental quotient >2 SDs below the population mean; blindness (visual acuity <6 of 60) or deafness (any hearing impairment requiring unimproved by amplification); and
- growth (weight, length, head circumference, mid-upper arm circumference, skinfold thickness) absolute change measured as grams or centimeters, standardized change measured as z-score, or percentile compared with a population reference.

The timing of the outcome assessment was at hospital discharge and at latest follow-up time recorded.

Search Methods

This is an update of a 2011 systematic review. Electronic databases were searched from January 1, 2011, to October 1, 2021. Databases included Medline (Ovid), Embase (Ovid), and Cochrane Central Register of Controlled Studies. In addition, we completed manual reference checks of existing reviews and of papers that were included in the review. All studies from the 2011 systematic review were included. Appendix 1 provides the search strategy used and Appendix 2 shows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart.

Selection of Studies and Data Extraction

Selection of studies and data extraction was conducted by 2 authors and followed standard methods. Data extracted included: country, study design, study setting (facility [infant born and followed-up in the facility until discharge] or “whole population” [ie, infant born...
and followed-up at home or the facility], gestational age, birth weight, and milks given in intervention and control groups.

**Assessment of Risk of Bias**

Two review authors judged the risk of bias using standard methods, including the ROBINS-I tool (risk of bias in nonrandomized studies of interventions). Where possible, funnel plots and Egger’s test were used to assess publication bias.

**Measurement of Treatment Effect**

For dichotomous data, we summarized results using risk ratios and, where this was not possible, odds ratios (ORs) with 95% confidence intervals (CIs). For continuous data, we summarized results using the mean difference (MD) with 95% CIs or standardized mean difference (SMD) when different methods or scales were used between studies.

We considered all studies to be highly heterogenous, so we used random-effects models to calculate pooled estimates for all outcomes. Where available, we used study level adjusted effect sizes to calculate pooled estimates and, when not available, we used raw data. We imputed missing data on the basis of Cochrane methods.

Restricted maximum likelihood estimates and Knapp–Hartung SEs were used. We also assessed forest plots visually for heterogeneity, and considered I-squared values >50% to represent substantial heterogeneity. All analyses were done using Stata 16.1.

**Subgroup and Sensitivity Analysis**

Our a priori subgroup analyses were:

1. gestational age and weight at birth (studies enrolling only infants <32 weeks’ gestation or <1.5 kg at birth compared with studies that did not restrict enrollment on the basis of gestational age or birth weight); and
2. exclusivity of mother’s own milk in the comparison arms (ie, studies providing mother’s own milk as the sole diet [ie, 100% mother’s own milk]) compared with studies providing other milks (eg, artificial infant formula, other animal milk, donor human milk) or foods (eg, porridge) mixed with mother’s own milk.

We had also planned to stratify analyses by high-, middle-, and low-income settings; however, there were no low-income studies and only 3 middle-income studies. We completed a sensitivity analysis to determine the robustness of the growth income by excluding 1 study that contributed a large sample size. We also planned a sensitivity analysis excluding studies at serious or critical risk of bias; however, only 1 study did not have serious or critical risk of bias.

**Summary of Findings and GRADE Table**

We prepared a summary of findings table for each outcome using Grading of Recommendations Assessment, Development and Evaluation (GRADE) and GRADEPro guideline development tool software to assess the quality of the body of evidence, consistency of effect, imprecision, indirectness, and publication bias for each outcome.

**RESULTS**

**Study Characteristics**

The search resulted in 5170 records. After screening titles and abstracts, 89 records were retrieved. Sixty-three reports were excluded with reasons (Appendix 2). We identified 26 new articles reporting on new studies. Combined with the 19 studies from the previous review, 42 studies were included in the narrative review and 36 studies provided data for meta-analysis.

Of the 42 studies there were no RCTs, but there were 34 cohort, 5 cross-sectional, and 3 case-control studies (Appendix 3). Studies were from 21 countries: Australia, Belgium, Chile, China, Germany, Ghana, Greece, Hong Kong, India, Israel, Italy, Japan, Nepal, Netherlands, New Zealand, Poland, Romania, Spain, Sweden, the United Kingdom, and the United States.

Thirty-six studies were implemented in NICUs and special care nurseries, and 6 were whole population studies.

In total, there were 89 638 preterm and/or LBW infants included in the review, of whom 74 656 were in the formula group and 14 982 were in the mother’s own milk group. Of the infants included in the review, 77 892 were infants born at <32 weeks’ gestational age and 76 796 were infants weighing <1.5 kg at birth. Twenty studies enrolled infants <32 weeks’ gestation and/or <1.5 kg at birth, and 22 studies did not restrict enrollment on the basis of gestational age or birth weight. In the intervention arm, 24 studies provided formula milk only (ie, formula was the sole diet) (Appendix 3 and 4). In the control arm, 9 studies provided mother’s own milk only (ie, mother’s own milk was the sole diet). The proportion of formula and mother’s own milk provided in each study is given in Appendix 4. All formula milks, where

---

†12,16,18,22,23,26,30–33,34–46,49–53, 57,59
‡17,19–21,25,27–29,34–40,47,48,54–56, 58,60
§12,16–18,21–23,26–29,32–35,37,38, 41–44,49,51,52,55,59,60

---
<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No. of Participants (Studies) Follow-up</th>
<th>Certainty of the Evidence (GRADE)</th>
<th>Relative Effect (95% CI)</th>
<th>Anticipated Absolute Effects</th>
<th>Anticipated Absolute Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neonatal mortality at latest follow-up: mean 116 d</td>
<td>9673 (5 observational studies)</td>
<td>⨁⨁⨁⨁ Very low(^A)</td>
<td>OR 1.26 (0.91–1.76)</td>
<td>57 per 1000</td>
<td>14 more per 1000 (5 fewer–39 more)</td>
</tr>
<tr>
<td>Necrotizing enterocolitis at latest follow-up: mean 44 d</td>
<td>3013 (15 observational studies)</td>
<td>⨁⨁⨁⨁ Low(^A)</td>
<td>OR 2.99 (1.75–5.11)</td>
<td>60 per 1000</td>
<td>100 more per 1000 (40 more–186 more)</td>
</tr>
<tr>
<td>Sepsis/infection at latest follow-up: mean 31 d</td>
<td>2572 (15 observational studies)</td>
<td>⨁⨁⨁⨁ Very low(^A)</td>
<td>OR 1.52 (0.98–2.37)</td>
<td>258 per 1000</td>
<td>88 more per 1000 (4 fewer–194 more)</td>
</tr>
<tr>
<td>Child cognitive development; assessed with validated child development assessment at follow-up (range: 91 wk–416 wk)</td>
<td>1560 (8 observational studies)</td>
<td>⨁⨁⨁⨁ Very low(^A)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Child language development; assessed with validated child development assessment at follow-up (range: 39 wk–104 wk)</td>
<td>587 (3 observational studies)</td>
<td>⨁⨁⨁⨁ Very low(^A)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Weight z-score (change from birth to discharge) follow-up: mean 52 d</td>
<td>74130 (6 observational studies)</td>
<td>⨁⨁⨁⨁ Very low(^C)</td>
<td>—</td>
<td>The mean growth: weight z-score (change from birth to discharge) ranged from 1.57 to –0.8 points</td>
<td>MD 0.03 points higher (0.15 lower–0.2 higher)</td>
</tr>
<tr>
<td>Weight z-score follow-up: range 39 wk–416 wk</td>
<td>271 (3 observational studies)</td>
<td>⨁⨁⨁⨁ Very low(^C)</td>
<td>—</td>
<td>The mean growth: weight z-score ranged from –1.31 to –0.5 points</td>
<td>MD 0.14 points higher (0.76 lower–1.05 higher)</td>
</tr>
<tr>
<td>Length (cm) at latest follow-up: mean 58 d</td>
<td>1048 (9 observational studies)</td>
<td>⨁⨁⨁⨁ Very low(^C)</td>
<td>—</td>
<td>The mean growth: length (cm) at latest follow-up ranged from 45.6 to 50.0 cm</td>
<td>MD 0.33 cm more (0.4 fewer–1.05 more)</td>
</tr>
<tr>
<td>Length z-score follow-up: range 39 wk–416 wk</td>
<td>271 (3 observational studies)</td>
<td>⨁⨁⨁⨁ Very low(^C)</td>
<td>—</td>
<td>The mean growth: length z-score ranged from –0.93 to –0.05 points</td>
<td>MD 0.06 points higher (0.81 lower–0.92 higher)</td>
</tr>
</tbody>
</table>
Information was provided, were artificial cow's milk protein-based. There were no other types of other protein base (eg, soy or goat milk protein) and no hydrolyzed formula in the studies. Seventeen studies provided donor milk in either the intervention or control groups. A total of 36 studies of 88,741 infants provided data for meta-analysis.

Risk of bias

A risk of bias assessment was completed for the 36 studies included in the meta-analysis. No studies had low risk of bias. Nineteen had critical, 16 had serious, and 1 had moderate risk of bias. Most biases were because of confounding (Appendix 5). No studies in the sepsis outcome had small study effects (ie, events ranging from 0 to 2) in the intervention and/or control groups (Egger test P = 0.0459; funnel plot 5.4). No other outcomes had obvious publication bias or small study effects in any analyses (funnel plots shown in Appendix 5).

Primary Outcomes

At discharge, there was little effect of the intervention (formula milk) on the primary outcome (all-cause mortality) (OR 1.26; 95% CI 0.91–1.76; I² = 0%; low certainty). No other outcomes had obvious publication bias or small study effects in any analyses (favored the intervention and/or control groups). No other outcomes had obvious publication bias or small study effects in any analyses.

Anticipated Absolute Effects

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>No. of Participants Follow-up</th>
<th>Certainty of the Evidence (GRADE)</th>
<th>Relative Effect (95% CI)</th>
<th>Risk With Any Human Milk</th>
<th>Risk Difference With Any Nonhuman Milk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head circumference (cm) follow-up: mean 45 d</td>
<td>1550 (3 observational studies)</td>
<td>△△△△ Very low ch</td>
<td>—</td>
<td>The mean growth: head circumference (cm) ranged from 30.9 to 34.5 cm</td>
<td>MD 0.26 cm higher (0.35 lower–0.87 higher)</td>
</tr>
</tbody>
</table>

The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). Patient or population: preterm or LBW infants (non-randomized studies of intervention); setting: during the neonatal period; intervention: any nonhuman milk; comparison: mother's own milk. GRADE working group grades of evidence: high certainty, we are very confident that the true effect lies close to that of the estimate of the effect; moderate certainty, we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different; low certainty, our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect; very low certainty: we have very little confidence in the estimate of the effect; true effect is likely to be substantially different from the estimate of the effect. NRSI, nonrandomized studies of intervention; — not applicable.

Risk of bias was downgraded by 2 levels: serious risk of bias (some studies did not account for confounding; it's possible classification of intervention status could have been affected by knowledge of the outcome or risk of the outcome). Downgraded by 1 level: serious inconsistency (moderate I² = 65%). Downgraded by 1 level of bias: serious risk of bias (no studies accounted for confounding). Downgraded by 1 level: serious inconsistency (high I² = 99%). Downgraded by 1 level: serious imprecision (small sample size). Downgraded by 1 level: serious inconsistency (small study population). Downgraded by 1 level: serious inconsistency (high I² = 80%). Downgraded by 1 level: serious inconsistency (high I² = 88%).
Subgroup and Sensitivity Analysis

There were no differences in the effect of the intervention on language outcomes in infants who received mother’s own milk as a sole diet compared with infants who received a mixture of mother’s own milk and other milks (Appendix 8).

A sensitivity analysis was completed for weight z-score at latest follow-up by excluding the study with the large sample size ($n = 72,997$). After removing this study, there was little change to results (MD 0.01; 95% CI –0.28 to 0.30; 5 studies; 1133 participants) (Appendix 9).

DISCUSSION

In our systematic review of 42 observational studies enrolling 89,638 preterm and LBW infants, we found that formula milks had little or no effect compared with mother’s own milk on mortality, severe infection, neurodevelopment, weight, length, or head circumference at discharge or latest follow-up, but found a threefold increase in necrotizing enterocolitis. We found no differential effect in the 17 studies enrolling only infants <32 weeks’ gestation or <1.5 kg at birth compared with the 19 studies that did not restrict enrollment on the basis of gestational age or birth weight. We also found no differential effect in the 8 studies providing mother’s own milk as the sole diet compared with the 28 studies providing other milks (eg, artificial infant formula, other animal milk, donor human milk) or foods (eg, porridge) mixed with mother’s own milk.

The previous systematic review of 19 observational studies enrolling 13,027 infants reported low certainty evidence that formula milk was associated with an increase in mortality and the combined outcome of severe infection and necrotizing enterocolitis compared with mother’s own milk in preterm and LBW infants. Formula milk was also associated with increased length and decreased neurodevelopmental outcomes, but no change in weight outcomes. Our search found 19 new observational studies and increased the number of participants contributing data to 88,741. The addition of new studies to the mortality (Svenningsen 1982, Ruys 2017), length (Costa-Orvay 2011, Madore 2017, Mol 2019, Pieltan 2001, Ruys 2017) and neurodevelopment (O’Connor 2003) outcomes did not substantially change the effects reported in 2011, though strength of effect for all outcomes was reduced and there was no improvement in the certainty or quality of the evidence.

Three Cochrane reviews assessed the effects of human milk in preterm and LBW infants. In 2019, a Cochrane review of RCTs of the effects of formula compared with mother’s own milk located no trials. Also in 2019, a Cochrane review of RCTs reported moderate certainty evidence that formula milk increased weight, length, and head growth, and had a higher risk of necrotizing enterocolitis compared with donor human milk. The trial data did not show an effect on all-cause mortality, or on long-term growth or neurodevelopment. A Cochrane
review of RCTs in 2020 also reported moderate certainty evidence that provision of multicomponent "fortifier" (powdered or liquid supplement with protein, carbohydrate, vitamins, and minerals added to human milk) increased short term in hospital weight, length, and head circumference compared with "unfortified" human milk, but evidence was insufficient to assess long-term effect on growth, neurodevelopment, mortality, or morbidity outcomes.61

In addition to the above review, 3 non-Cochrane reviews were identified.63–65 Each of these reviews only investigated 1 outcome per review, which included a range of in-hospital growth outcomes, bronchopulmonary dysplasia, and necrotizing enterocolitis. One review found that, for preterm and LBW infants, there was inconclusive evidence on the effect of formula compared with exclusive human milk on growth parameters including change in weight z-scores and head circumference.65 For preterm infants receiving exclusive human milk, there was an improvement in bronchopulmonary dysplasia compared with exclusive formula-fed infants; however, the review combined both RCTs and observational studies in their analysis.64 Similar to our observational study review, preterm infants who received human milk had a reduced risk of developing necrotizing enterocolitis than those who received formula.63 However, unlike our review, these reviews did not differentiate between whether infants received mainly mother's own milk; only that they received exclusive (ie, mother's own milk and/or donor milk) or any human milk.

Our review had some limitations. All evidence was low to very low certainty because of problems with confounding bias, unexplained heterogeneity, small sample sizes, and imprecision in many studies. Thirty-six studies recruited infants from health facilities only; there were only 6 studies that recruited infants from the whole population. Twenty-two studies recruited infants >32 weeks' gestation. We had also planned to stratify analyses by high-, middle-, and low-income settings; however, there were no low-income studies and only 3 middle-income studies. We also planned a sensitivity analysis excluding studies at serious or critical risk of bias; however, only 1 study did not have serious or critical risk of bias. One study contributed a large sample size (n = 72 997) to the growth outcome and had the potential to bias outcomes. However, when we removed this study from the analysis, there was little change to the results. There is also much potential for miscategorization between intervention and control groups (ie, switching between formula and mother's own milk groups). This is because of the common practice of providing formula if mother's own milk is not available and active promotion of breastfeeding in formula-fed infants. However, in our study, we defined the intervention and control groups as those receiving >50% (ie, "most or majority") formula (intervention) or mother's own milk (control) over the entire study period. We also assessed sole or exclusive diet as a subgroup. Some studies did not describe the amount of formula or mother's own milk provided but we were able to estimate most or majority >50% for all studies. Other strengths of our study were the comprehensive search strategy and the inclusion of all study designs.

Overall, in preterm and LBW infants, our review shows low- to very low-certainty evidence that feeding with formula milks compared with mother's own milk, either as a sole diet or mixed with other milks, has little effect on all-cause mortality, severe infection, growth, or neurodevelopment, and a higher risk of developing necrotizing enterocolitis. However, the quality of the observational evidence base must be improved. It is especially important to control studies for confounding and to conduct more studies in low-income countries and in infants born and cared for outside of health facilities.

**ABBREVIATIONS**

CI: confidence interval
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
LBW: low birth weight
MD: mean difference
OR: odds ratio
RCT: randomized controlled trial
SMD: standardized mean difference

**REFERENCES**


